Precision Medicine
The Impact of Companion Diagnostic Device Measurement Performance on Clinical Trial Designs

Meijuan Li, Ph.D.
Acting Deputy Director, Division of Biostatistics/CBER
Chief, Diagnostic Statistics Branch II, Division of Biostatistics/CDRH
U.S. Food and Drug Administration

This talk is based on the following publication/work completed at CDRH/FDA

- *Li, M., Yu, T., Hu, Y., The Impact of Companion Diagnostic Device Measurement Performance on Clinical Validation of Personalized Medicine. Statistics in Medicine*
- *Li, M. Statistical consideration and challenges in bridging study of personalized medicine. Journal of Biopharmaceutical Statistics.*
- *Li, M., Pennello, G., Wu, J., et al. Personalized Medicine. Encyclopedia of Biopharmaceutical Statistics*

Uses of Biomarkers

- **Diagnosis,** in symptomatic patients
- **Early detection (screening)**, enabling intervention at an earlier and potentially more curable stage than under usual clinical diagnostic conditions

Precision Medicine

- **Risk assessment,** leading to preventive interventions for those at sufficient risk
- **Prognosis**, allowing for more (less) aggressive therapy for patients with worse (better) prognosis
- **Monitoring of disease,** response during therapy, with potential for adjusting level of intervention (e.g. dose) on a dynamic and personal basis
- **Predictive,** E.g., predicts safety, efficacy, or PK/PD of a specific therapy, thereby providing guidance in selecting it for patients or tailoring its dose (Companion or Complementary Diagnostic)

Diagnostic Devices

- **Diagnostic device:** A diagnostic medical device is an instrument, apparatus, implement, machine, contrivance, implant, in vitro reagent, or other similar or related article, including a component part, or accessory which is recognized in the official National Formulary, the United States Pharmacopoeia, or any supplement to them, intended for use in the diagnosis of disease or other conditions
- Biomarker is measured or detected by a diagnostic device
  - **Device ≠ biomarker**
  - A biomarker is only as good as the device used to measure/detect it

Companion Diagnostic Devices

Companion Diagnostic Device

- A companion diagnostic device (or test) is an in vitro diagnostic device that provides information ***that is essential*** for the safe and effective use of a corresponding therapeutic product (*In Vitro Companion Diagnostic Devices FDA GUIDANCE, 2014*)

- Companion diagnostic devices (CDx) are instruments, systems, apparatus, or platforms used to detect /measure ***predictive biomarker*** used in personalized medicine

Companion Diagnostic Device

- Regulatory review and approval/clearance
  - if a new drug needs a CoDx, it may not be approved without co-approval of an in vitro diagnostic (IVD) or other test for this purpose.

CDx Intended Uses

- **Predictive tests**: CDx predicts a differential effect of the treatment on an clinical outcome e.g.
  - identify patients who are most likely to benefit from a particular therapeutic product
  - identify patients who are likely to be at increased risk for serious adverse reactions as a result of treatment with a particular therapeutic product
  - need to demonstrate clinically and statistically significant treatment by CDx interaction

- **Selection tests:** CDx selects patients for drug treatment
  - the clinical trial may be NOT designed to predict a differential effect of the treatment on an clinical outcome (i.e. treatment efficacy on marker negative group defined by CDx is unknown)
  - need to demonstrate clinically and statistically significant treatment efficacy in selected patient group defined by CDx

Regulatory Review of a Companion Diagnostic Device

- Two types of CDx performance which FDA review focuses on
  - **Measurement performance**: parameters associated with the measuring capability of a CDx
  - **Clinical performance**: parameters associated with the capability of a CDx for predicting treatment responses or selecting patients for the treatment

CDx Measurement Performance

- **Pre-analytical:** related to specimen collection (including timing, technique (aliquoting, pipetting, and retrieval), and processing), as well as handling and storage (including time, temperature, humidity, and volume)
- **Analytical:** related to the precision and accuracy of the test method and factors which may interfere with a particular assay
- **Post-analytical:** related to data entry and calculations by laboratory staff, result validation, interpretation of the result, data transfer and the method used to report the results (electronic, paper or telephone)

Analytical Performance Studies

- Analytical accuracy
- Precision (repeatability, intermediate precision, and reproducibility)
- Limits of detection (LoD), limit of quantitation (LoQ), limit of blank (LoB)
- Reference ranges
- Linearity
- Matrix effect
- Stability (of the test system and the measurand)
- Interferences
- Carry-over
- ……..

Analytical Accuracy

- **Analytical accuracy** measures how well the diagnostic device output agrees with true marker status

- Denote G be the true marker unknown status.
  Let G=1 and G=0 for true unknown marker positive and negative, respectively

- G is measured by an imperfect companion diagnostic device M.
  Let M=1 and M=0 for measured marker positive and negative, respectively

Analytical Accuracy Measures

- Analytical accuracy of device M can be quantified by
  - **PPA and NPA pair**
    - Positive percentage agreement (PPA) defined as Pr(M=1|G=1)
    - Negative percentage agreement (NPA) defined as Pr(M=0|G=0)
  - **PPV and NPV pair**
    - Positive predictive value (PPv) defined as Pr(G=1|M=1)
    - Negative predicative value (NPv) defined as Pr(G=0|M=0)
    - Note PPV and NPV are marker positivity dependent where positivity of G is $p = \Pr(G = 1)$ in the intended use population of M
- Two different types of misclassification for M
  - When G=1, M=0 (false negative)
  - When G=0, M=1 (false positive)

(PPA and NPA) vs (PPV and NPV)

- $p = Pr(G = 1)$
- $pm = Pr(M = 1) = pPPA + (1 - p)NPA = \frac{pNPv-1}{NPv+PPv-1}$
- $PPv = \frac{pPPA}{pPPA+(1-p)NPA}$ & $NPv = \frac{(1-p)NPA}{(1-p)NPA+p(1-PPA)}$
- $PPA = \frac{pm\ PPv}{pmPPv+(1-pm)NPv}$ & $NPA = \frac{(1-pm)NPv}{(1-pm)NPv+pm(1-PPv)}$

Precision

- **Measurement precision:** the closeness of agreement between replicate measurements on the same object (e.g., sample) under specified testing conditions
  - repeatability conditions (replicate measurements on the same or similar objects under the same or similar conditions)
  - reproducibility conditions (replicate measurements on the same or similar objects using different operating conditions)
  - some other sets of intermediate conditions

Example: continuous measurement repeatability study:

1. Samples: patient plasma samples
2. Repeatability study design
   - One run per day over *m non-consecutive* days
   - Two replicates per run
   - The within-run precision (or repeatability) standard deviation can be calculated as

$$\widehat{sd} = \sqrt{\sum_{i=1}^{m} (x_{1i} - x_{2i})^2 / 2m}$$

where: m= total number of days, x_{1i} and x_{2i} are result for replicates 1 and 2 on day i, respectively.

$$CV\% = 100 * sd/\mu, \quad \widehat{CV}\% = 100 * \widehat{sd}/\bar{x}$$

Confirmatory Clinical Trial for CDx and the Corresponding Therapeutic Product

Trial Assumptions

- Patients in the clinical trial are randomized to either therapeutic treatment arm (denoted as ***t=1)*** or control arm (denoted as ***t=0)*** with a ratio of 1:1 stratified by *M* or only enroll M+ patients (enrichment design)

- Assume continuously valued clinical outcome **Z** only depends on *G* and T.

- Assume that the true marker status G does not depend on treatment T, and denote $G = 1, 0$ for true marker positive and negative, respectively

- G is measured by an imperfect companion diagnostic device M. M=1 and M=0 for marker positive and negative, respectively

Notation - Device Measurement Outcome

- Denote device underlying continuous signal (measurement) ***Y*** are
  - G+ patients: $Y_1 \sim N(\vartheta_1, {w_1}^2 + {s_1}^2)$,
  - G- patients: $Y_0 \sim N(\vartheta_0, {w_0}^2 + {s_0}^2)$

  where $s = (s_1, s_0)$ be device measurment imprecision for G+ and G- patients, respectively; ${w_1}^2, {w_0}^2$ are biological variability for G+ and G-, respectively.

- Denote c is the cutoff, i.e. $y_1 \ (or\ y_0) \geq c,\ M = 1$ and 0 otherwise

Notation - Device Measurement Outcome

- Analytical Accuracy when the device precision is $s = (s_1, s_0)$
  - $PPA_s = \int_c^{\infty} \frac{1}{\sqrt{2\pi(w_1{}^2+s_1{}^2)}} e^{-\frac{(y_1-\vartheta_1)^2}{2(w_1{}^2+s_1{}^2)}} dy_1$; $NPA_s = \int_{-\infty}^{c} \frac{1}{\sqrt{2\pi(w_1{}^2+s_1{}^2)}} e^{-\frac{(y_0-\vartheta_0)^2}{2(w_1{}^2+s_1{}^2)}} dy_0$
  - $\boldsymbol{PPv_s = pPPAs/\{pPPAs + (1-p)(1-NPAs)\}}$
  - $\boldsymbol{NPv_s = (1-p)NPAs/\{(1-p)NPA_s + p(1-PPAs)\}}$
  - $p = \Pr(G = 1)$

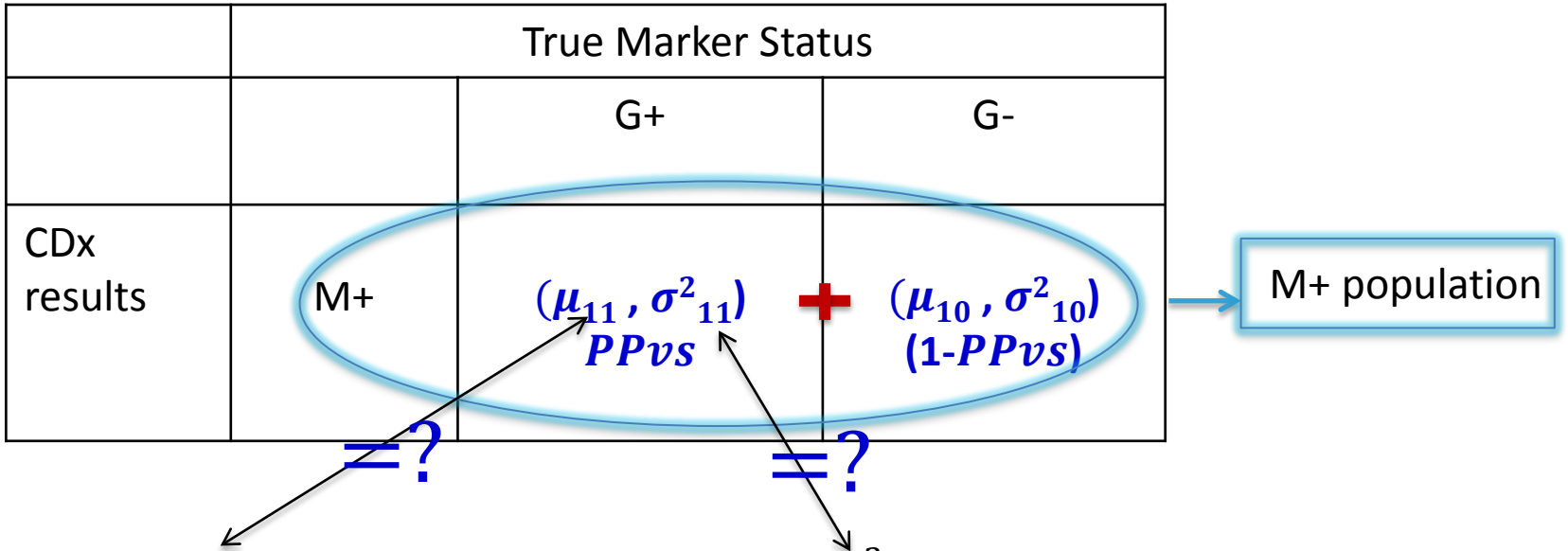
New Treatment Arm: t=1

		True Marker Status	
		G+ Population	G- Population
CDx results	M+ population	$(\mu_{11}, \sigma^2_{11})$	$(\mu_{10}, \sigma^2_{10})$
	M-population	$(\mu_{11}, \sigma^2_{11})$	$(\mu_{10}, \sigma^2_{10})$

(Mean, and Variance)|t=1, **G+**

(Mean, and Variance)|t=1, **G-**

New Treatment Arm: t=1



(mean, and Variance)|t=1,M+ = (θ_{11}, v^2_{11})

$\theta_{11} = PPvs\ \boldsymbol{\mu_{11}} + (1 - PPvs)\mu_{10}$

$v^2_{11} = PPvs\ \boldsymbol{\sigma_{11}}^2 + (1 - PPvs)\sigma_{10}{}^2 + PPvs_{(1 - PPvs)}(\mu_{11} - \mu_{10})^2$

New Treatment Arm: t=1

	True Marker Status		
		G+	G-
CDx results	M-	$(\mu_{11}, \sigma^2_{11})$ (1-$NPvs$)	$(\mu_{10}, \sigma^2_{10})$ $NPvs$

→ M- population

(mean, and Variance)|t=1,M- = (θ_{10}, v^2_{10})

$\theta_{10} = (1 - NPvs)\mu_{11} + NPvs\mu_{10}$

$v^2_{10} = (1 - NPvs)\sigma_{11}^2 + NPvs\sigma_{10}^2 + NPvs_{(1-NPvs)}(\mu_{11} - \mu_{10})^2$

Control Arm: t=0

	True Marker Status		
		G+	G-
CDx results	M+	$(\mu_{01}, \sigma^2_{01})$ $PPvs$ +	$(\mu_{00}, \sigma^2_{00})$ (1-$PPvs$)

→ M+ population

(mean, and Variance)|t=0,M+ = (θ_{01}, v^2_{01})

$\theta_{01} = PPvs\mu_{01} + (1 - PPvs)\mu_{00}$

$v^2_{01} = PPvs\sigma_{01}{}^2 + (1 - PPvs)\sigma_{00}{}^2 + PPvs_{(1 - PPvs)}(\mu_{01} - \mu_{00})^2$

Control Arm: t=0

	True Marker Status		
		G+	G-
CDx results	M-	$(\mu_{01}, \sigma^2_{01})$ (1-$NPvs$)	$(\mu_{00}, \sigma^2_{00})$ $NPvs$

→ M- population

(mean, and Variance)|t=0,M- = (θ_{00}, v^2_{00})

$$\theta_{00} = (1 - NPvs)\mu_{01} + NPvs\mu_{00}$$

$$v^2_{00} = (1 - NPvs)\sigma_{01}{}^2 + NPvs\sigma_{00}{}^2 + NPvs_{(1 - NPvs)}(\mu_{01} - \mu_{00})^2$$

Notation

- Denote $\mu_{11}, \mu_{10}, \mu_{01}, \mu_{00}$ be the drug efficacies for patients in $t = 1$ & $G = 1, t = 1$ & $G = 0, t = 0$ & $G = 1, t = 0$ & $G = 0$, respectively.

- Denote $\theta_{11}, \theta_{10}, \theta_{01}, \theta_{00}$ be the mean drug efficacies for patients in $t = 1$ & $M = 1, t = 1$ & $M = 0, t = 0$ & $M = 1, t = 0$ & $M = 0$, respectively.
  - $\theta_{11} = PPvs\mu_{11} + (1 - PPvs)\mu_{10}$, $\theta_{01} = PPvs\mu_{01} + (1 - PPvs)\mu_{00}$
  - $\theta_{10} = (1 - NPvs)\mu_{11} + NPvs\mu_{10}$, $\theta_{00} = (1 - NPvs)\mu_{01} + NPvs\mu_{00}$

- Denote ${\sigma^2}_{11}, {\sigma^2}_{10}, {\sigma^2}_{01}, {\sigma^2}_{00}$ be the variance of clinical outcome for patients in $t = 1$ & $G = 1, t = 1$ & $G = 0, t = 0$ & $G = 1, t = 0$ & $G = 0$, respectively.

- Denote ${v^2}_{11}, {v^2}_{10}, {v^2}_{01}, {v^2}_{00}$ be the variance of clinical outcome for patients in $t = 1$ & $M = 1, t = 1$ & $M = 0, t = 0$ & $M = 1, t = 0$ & $M = 0$, respectively.
  - ${v^2}_{11} = PPvs{\sigma_{11}}^2 + (1 - PPvs){\sigma_{10}}^2 + PPvs_{(1 - PPvs)}(\mu_{11} - \mu_{10})^2$
  - ${v^2}_{10} = (1 - NPvs){\sigma_{11}}^2 + NPvs{\sigma_{10}}^2 + NPvs_{(1 - NPvs)}(\mu_{11} - \mu_{10})^2$
  - ${v^2}_{01} = PPvs{\sigma_{01}}^2 + (1 - PPvs){\sigma_{00}}^2 + PPvs_{(1 - PPvs)}(\mu_{01} - \mu_{00})^2$
  - ${v^2}_{00} = (1 - NPvs){\sigma_{01}}^2 + NPvs{\sigma_{00}}^2 + NPvs_{(1 - NPvs)}(\mu_{01} - \mu_{00})^2$

Notation

- Mean outcome difference between treatment arms
  - G+ patients: $\mu_1 = \mu_{11} - \mu_{01}$
  - G- patients: $\mu_0 = \mu_{10} - \mu_{00}$
  - M+ patients: $\theta_1 = (\theta_{11} - \theta_{01}) = PPvs\mu_1 + (1 - PPvs)\mu_0$
  - M- patients: $\theta_0 = (\theta_{11} - \theta_{01}) = (1 - NPvs)\mu_1 + NPv_s\mu_0$

- Variance of outcome
  - G+ patients: $\sigma^2{}_1 = \sigma^2{}_{11} + \sigma^2{}_{01}$
  - G- patients: $\sigma^2{}_0 = \sigma^2{}_{10} + \sigma^2{}_{00}$
  - M+ patients: $V^2{}_1 = PPvs\sigma^2{}_1 + (1 - PPvs)\sigma^2{}_0 + PPvs(1 - PPvs)\tau^2$
  - M- patients: $V^2{}_0 = (1 - NPvs)\sigma^2{}_1 + NPvs\sigma^2{}_0 + NPvs(1 - NPvs)\tau^2$

  where $\tau^2 = (\mu_{11} - \mu_{10})^2 + (\mu_{01} - \mu_{00})^2$

Selection Test: Impact on Treatment Effects

- The treatment effect in G+ patients: $\mu_1 = \mu_{11} - \mu_{01}$
- The treatment effect in M+ patients: $\theta_1 = (\theta_{11} - \theta_{01}) = PPvs\mu_1 + (1 - PPvs)\mu_0$

- The proportion of dilution in treatment effect in M+ is

$$pd = \frac{\mu_1 - \theta_1}{\mu_1} = (1 - PPvs)\left(1 - \frac{\mu_0}{\mu_1}\right) = (1 - PPvs)(1 - r)$$

- pd is a function of both $PPvs$ and r
- For a fixed value of $PPvs$, pd is negatively related to *r* e.g. when *r* decreases by 1% for any $r \in [0, 1)$, pd will increase by $(1 - PPvs)/100$
- For a fixed value of *r*, pd is also negatively related to $PPvs$

Selection Test: Impact on Variances

- Variance for G+ patients: $\sigma^2{}_1 = \sigma^2{}_{11} + \sigma^2{}_{01}$
- Variance for M+ patients : $V^2{}_1$
- Let $f(PPvs) = V^2{}_1 = PPvs\sigma^2{}_1 + (1 - PPvs)\sigma^2{}_0 + PPvs(1 - PPvs)\tau^2$

  where $f(PPvs)$ is a differentiable function
  - $f'(PPvs)$ is monotonically decreasing on its domain of $PPvs \in (p, 1],\ p = \Pr(G+)$
  - $f''(PPvs) = -2\tau^2 < 0$
  - $f(PPvs)$ is strictly concave on its domain of $PPvs \in (p, 1]$
- Let $k = [0.5 + (\sigma_1{}^2 - \sigma_0{}^2)/2\tau^2]$
  - If $k > 1, V^2{}_1$ increases for all $PPvs \in (p, 1]$
  - If $k < p, V^2{}_1$ decreases for all $PPvs \in (p, 1]$
  - If $k \in (p, 1], V^2{}_1$ increases for all $PPvs \in (p, k]$ and decrease otherwise

Selection Test: Impact on Sample Size

- Primary hypothesis testing: we are interested whether the drug efficacy θ_1 is significantly different from 0 or not:

$$\text{H0: } \theta_1 \leq 0 \qquad vs. \qquad \text{Ha: } \theta_1 > 0$$

  - The sample size required in the clinical trial for the treatment effect in M+ group is asymptotically given

$$n_1 = c_{\alpha\beta} \frac{\sigma^2_0 + PPvs(\sigma^2_1 - \sigma^2_0) + PPvs(1 - PPvs)\tau^2}{[PPvs(\mu_1 - \mu_0) + \mu_0]^2}$$

$$\text{where } c_{\alpha\beta} = 2(Z_{1\_\alpha} + Z_{1\_\beta})^2 \qquad \tau^2 = (\mu_{11} - \mu_{10})^2 + (\mu_{01} - \mu_{00})^2$$

Property of function $g(PPvs)$

- Let $g(PPvs) = n_1 = c_{\alpha\beta}\frac{\sigma^2_0 + PPvs(\sigma^2_1 - \sigma^2_0) + PPvs(1-PPvs)\tau^2}{[PPvs(\mu_1 - \mu_0) + \mu_0]^2}$

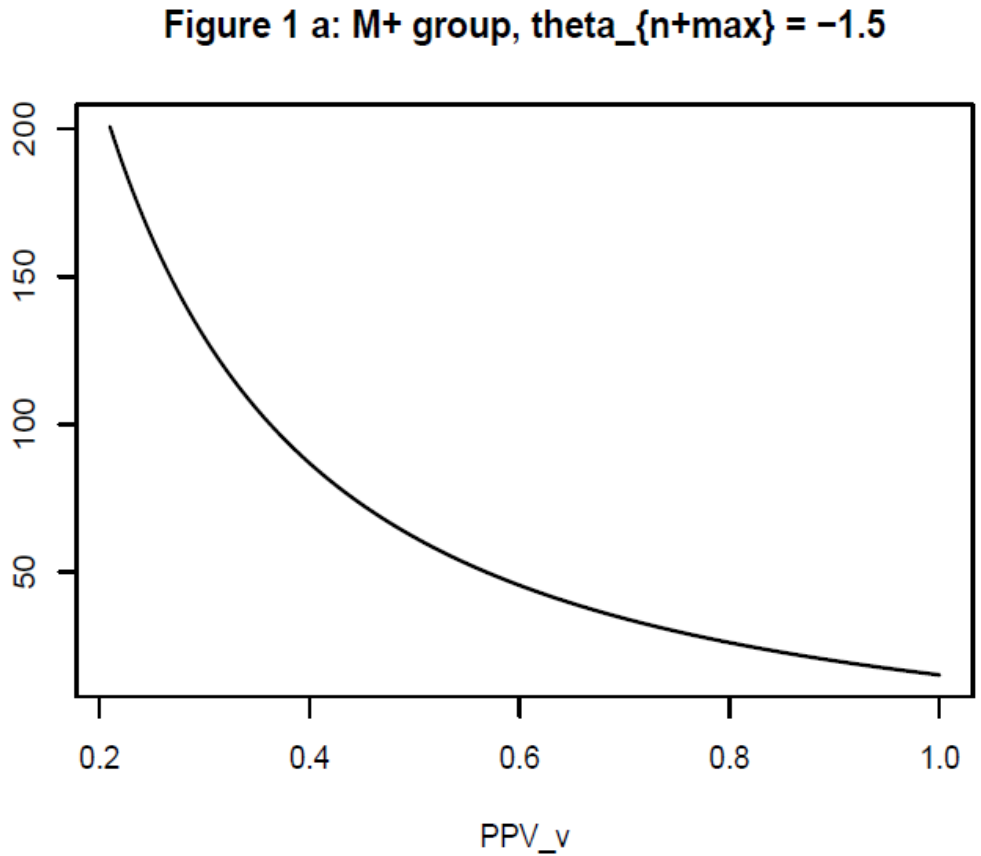
  $- \ g' = c_{\alpha\beta}\frac{(\sigma^2_0 + \sigma^2_1 + \tau^2)\mu_0 - 2\sigma^2_0\mu_1 + PPvs\{(\sigma^2_1 - \sigma^2_0)(\mu_1 - \mu_0) - \tau^2(\mu_1 + \mu_0)\}}{[PPvs(\mu_1 + \mu_0) + \mu_0]^3}$

  - the denominator of the first derivative is a strict increase function of $PPvs$
  - the numerator guarantees that the sign of g' can change for at most once on the whole real axis
  - The solution to the equation asserts that the extreme value of n_1 occurs when

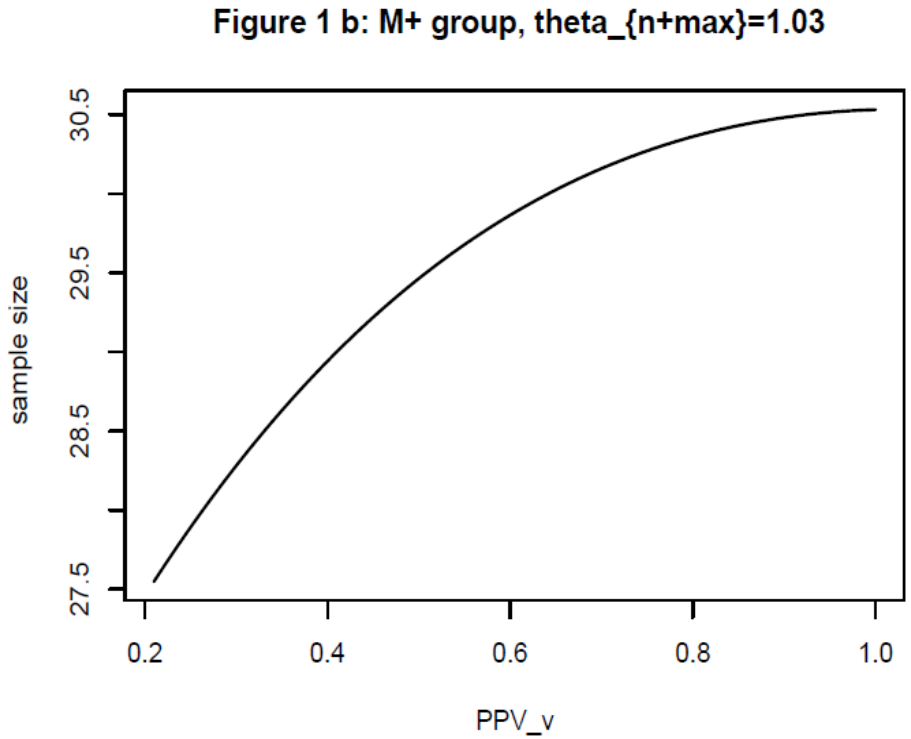
$$PPvs = \theta_{n+max} = \frac{2\sigma^2_0\mu_1 - (\sigma^2_1 + \sigma^2_0 + \tau^2)\mu_0}{[(\sigma^2_1 - \sigma^2_0)(\mu_1 - \mu_0) - \tau^2(\mu_1 + \mu_0)]}$$

- Impact of the device precision and accuracy on sample size n_1 is complex
  - Denote $\theta_{n+max} = \frac{2\sigma^2_0\mu_1 - (\sigma^2_1 + \sigma^2_0 + \tau^2)\mu_0}{[(\sigma^2_1 - \sigma^2_0)(\mu_1 - \mu_0) - \tau^2(\mu_1 + \mu_0)]}$
  - If $\theta_{n+max} < p$, n_1 decrease as PPv_s increases for all $PPv_s \in (p, 1]$
  - if $\theta_{n+max} \geq 1$, n_1 increase as PPv_s increases for all $PPv_s \in (p, 1]$ → **very unlikely!**
  - if $p \leq \theta_{n+max} < 1$, n_1 increase as PPv_s increases for all $PPv_s \in (p, \theta_{n+max}]$ and n_1 decreases as PPv_s increases for all $PPv_s \in (\theta_{n+max}, 1]$

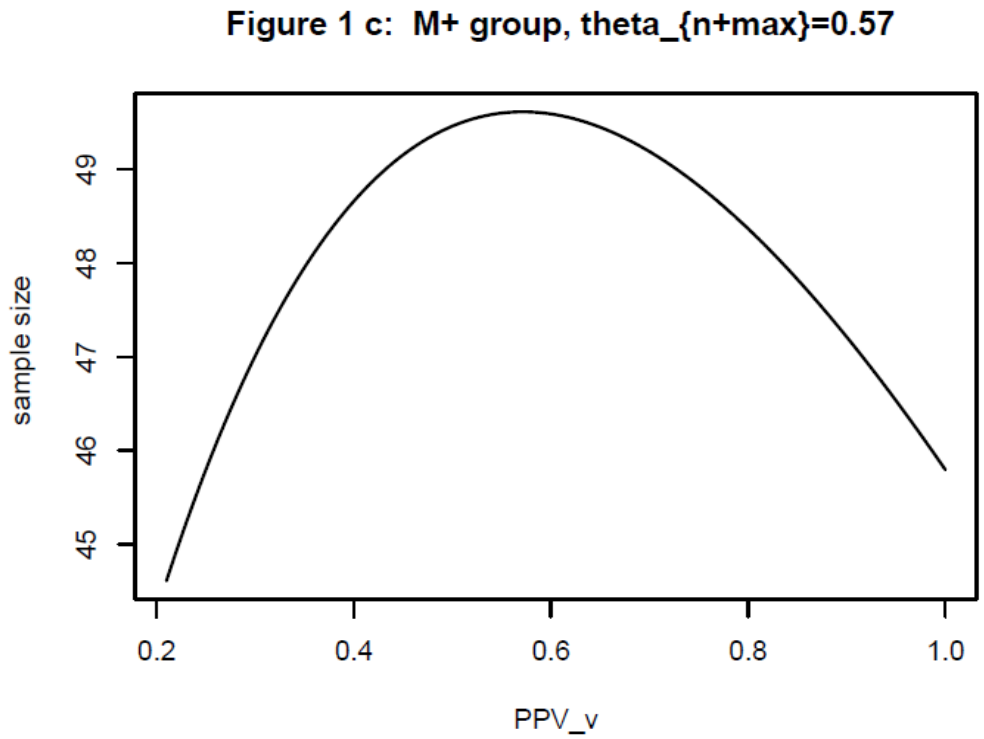
Impact on Sample Size



Impact on Sample Size



Impact on Sample Size



Predictive Test: Impact on Sample Size

- Primary hypothesis testing:

$$\text{H0: } (\theta_1 - \theta_0) \leq 0 \quad vs. \quad \text{Ha: } (\theta_1 - \theta_0) > 0$$

  - M+ patients: $\theta_1 = (\theta_{11} - \theta_{01}) = PPvs\mu_1 + (1 - PPvs)\mu_0$
  - M- patients: $\theta_0 = (\theta_{11} - \theta_{01}) = (1 - NPvs)\mu_1 + NPv_s\mu_0$

- The sample size required in clinical trial

$$\boldsymbol{n = c_{\alpha\beta}\ \frac{V_1^{\,2}/Pr(M=1) + V_0^{\,2}/Pr(M=0)}{[(PPv_s + NPvs - 1)(\mu_1 - \mu_0)]^2} = g(PPv_s, NPv_s)}$$

where

$$\Pr(M = 1) = \frac{p + NPVs - 1}{PPV_s + NPVs - 1},\ \Pr(M = 0) = (1 - P(M = 1)$$

$$V^2{}_1 = PPvs\sigma^2{}_1 + (1 - PPvs)\sigma^2{}_0 + PPvs(1 - PPvs)\tau^2$$

$$V^2{}_0 = (1 - NPvs)\sigma^2{}_1 + NPvs\sigma^2{}_0 + NPvs(1 - NPvs)\tau^2$$

Property of function $g(PPv_s, NPv_s)$

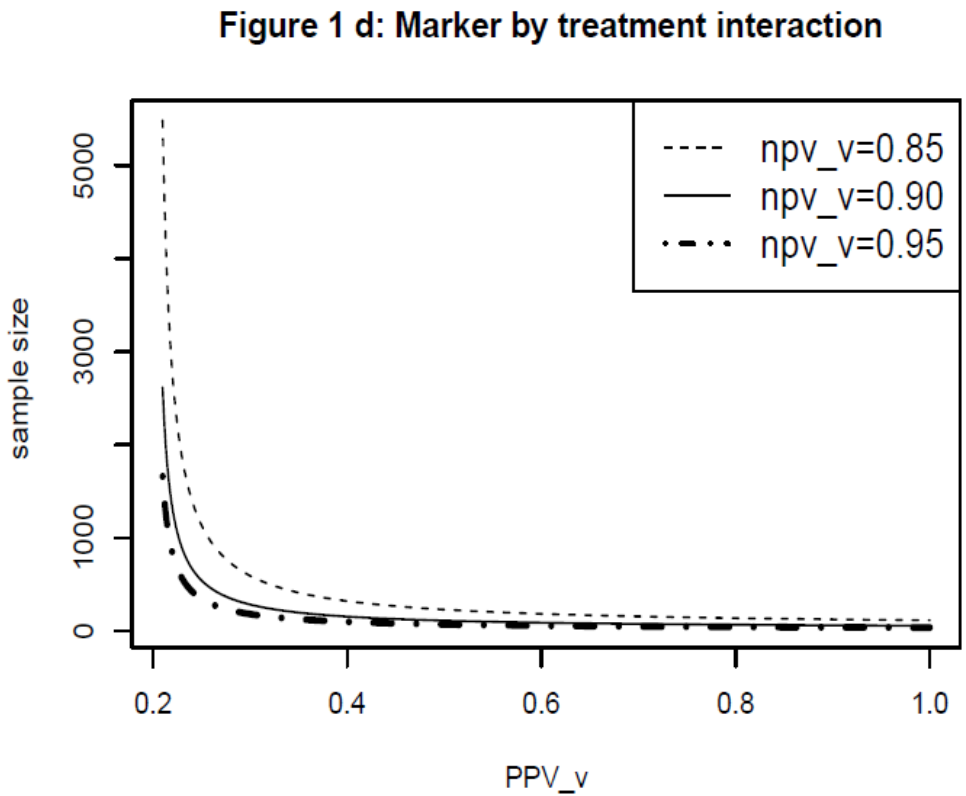
- For a fixed value of PPv_s, n is a strict increasing function of $NPvs$ on its domain.
- For a fixed value of NPv_s, n is a strict increasing function of $PPvs$ on its domain.
- It is sufficient to prove the monotonicity of n with respect to PPv_s and it is easy to transfer the property to NPv_s

Property of function $g(PPv_s, NPv_s)$

- **Proof**
  - $n \propto \frac{\sigma_0^2+(\sigma_1^2-\sigma_0^2+\tau^2)PPv_s-\tau^2PPv_s}{(p+NPvs-1)(PPvs+NPvs-1)}+\frac{\sigma_1^2+(\sigma_0^2-\sigma_1^2+\tau^2)NPv_s-\tau^2NPv_s}{(PPvs-p)(PPvs+NPvs-1)}$
    $= g(PPv_s)$
  - $\frac{\partial g(PPvs)}{\partial PPvs} \propto \frac{-(\sigma^2_1-\sigma^2_0+\tau^2)(1-NPvs)-\sigma^2_0+2\tau^2(1-NPvs)PPv_s-\tau^2PPv_s^2}{(PPvs+NPvs-1)2}$

  - The sign of the fraction only depends on the numerator, which is a quadratic function of PPv_s. Note that the extreme point of the numerator occurs with a maximum value $-\tau^2NPv_s(1-NPvs)-\sigma^2{}_0NPv_s-\sigma^2{}_{1(}1-NPvs) < 0$
  - Conclude that $\frac{\partial g(PPvs)}{\partial PPvs}$<0 for all $(PPv_s + NPvs) > 1$
  - It implies that n is decreasing with PPv_s because both terms of $g(PPv_s)$ are decreasing when PPv_s increases.

Sample Size (n) vs (PPv_s, NPv_s)



Conclusion

- In personalized medicine, CDx measurement performance impacts the clinical trial design including
  - Attenuation of the treatment effect
  - Influence on the variances of clinical outcome
  - Influence on required larger sample size and the sample size is bounded below when the device has perfect accuracy and precision.
  - The impact of CDx measurement performance on the clinical trial design is complicated
- CDx $\longleftrightarrow$ marker $\longleftrightarrow$ drug
  - Choose good marker (ALK, BCR/ABL gene translocation/rearrangement) which is highly liked to clinical outcome,
  - Choose a good CDx with high measurement performance
  - e.g. using a well-characterized companion diagnostic
- Future research: methods for time-to-event data